Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

The Importance of Data Policy and Governance for a Citizens Project using ML in Ethiopia

Ali Endris Mohammed

P.h.D Student, Data Science Lab

Department of Computer Science and Engineering (CSE) – SoEEC
Adama Science and Technology University

Joint ICTP-IAEA Advanced School/workshop on Machine
Learning in Citizen Science
3rd March 2023

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Outline

Introduction

Motivation

Data access and sharing in Ethiopia

Data Management policy

Conclusion

Reference

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Figure: Adama Science and Technology University (ASTU)

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Introduction

Recently, Machine Learning(ML) has played a significant role in enhancing business decision-making for different sectors. -Example of service

- From recommendation System to personalize learning, precision medicine, and precision agriculture.
- From Self-driving autonomous car to UAV
- From simulation to Metaverse, Digital Twins.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Introduction ...

Data is the backbone of this AI/ML approach, and its model is data-intensive for training and learning process.
-With the advancements of IoT, Wireless technologies, mobile applications, etc. the source of data is everywhere. Such as:-

- Industries, Academic institutions,
- Government service.
- From public

Nowadays, data become a strategic asset of any institution.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Motivation

Three universities ASTU, AASTU, and AAU are included in the case studies.

- From 2015-2020 more than 50 research and community service project proposals were accepted in the RTT office.
- Among them only 25 percent completed the remaining not completed yet.

The major reason mentioned in their report is the challenges of getting a dataset.

- Did we have a platform to accept the data from the source?
- Are we ready to utilize the data generated from different stakeholders?

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Data access and sharing in Ethiopia

Researchers face complex challenges to get data. Why it is very difficult?

- Getting the right formats of data
- Lack of awareness about the importance of data
- Absence of data management policy and guidelines

Traffic Accident data

·l·/·l·	Driver Age	The damage caused by the accident				Total
		Death	Serious Injuries	Minor injuries	Property damage	
1	< 18	1	3	8	38	50
2	Between 18 - 30	126	213	274	1677	2290
3	Between 31-50	116	220	219	1777	2332
4	>51	24	84	76	839	1023
5	Unknown	51	106	75	358	590
Total		318	626	652	4689	6285

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

The Challenges on data collection and management

Still, the following challenge exists.

- Infrastructure
- Lack of awareness among stakeholders
- Absence of public data repositories.
- Lack of collaboration between stakeholders.
- Government policy

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Efforts to deal with the problem

- In 2019 the Ethiopian public health institute (EPHI) tries to design a data repository for researchers with access and sharing guidelines.
- However, the scope is limited to the data source from:-
  - Laboratory health data
  - From RTT, PHEM, and
  - From national laboratory data.

-A framework for data collection and data format not included in the guideline

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Efforts to deal with the problem ...

Ethiopian Education and Research Network (EthERNet) was
established in 2001 EC to build the capacity of public universities
to share educational resources and research among member
institutions locally and globally.

- EthERNet linked with schoolNet and WoredaNet to extend its
  scope.
- Provides a storage and computational platform for researchers.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

EtherNet

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Data Management policy

Besides that, this huge amount of data with heterogeneity and
high speed needs a comprehensive data management policy in
terms of data collection, access, sharing, storage, and participation
of the respected stakeholders to benefit from it.

Why do we need a Data management policy?

- To facilitate data sharing and access for the researcher
- To manage data quality
- To control standards and ethics
- To ensure citizens' privacy and responsibilities.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Citizen Science

Participatory Data Management Framework
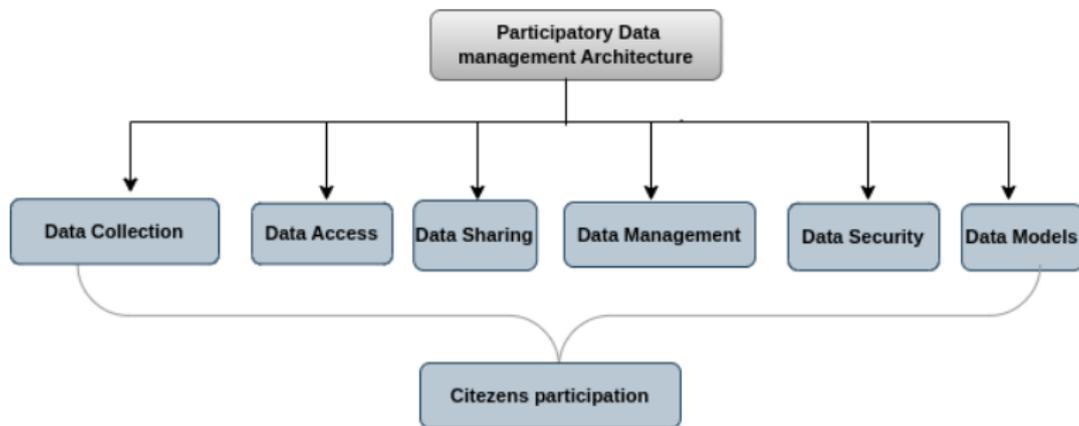
Figure: Data Management Model

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Citizen Science

Mainly, designing the framework with the following procedures
need to be in place

- Creating awareness
- Preparing data collection template
- Recording raw data
- Store in the repository.
- Setting access and sharing policies to ensure privacy

This is where citizen science brings the full involvement:-

- Design and planning
- Data collection
- Data analysis
- Dissemination and actions

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Data policy and Guideline for Citizen Science

- Protecting citizen privacy: Citizens have the right to privacy, and it's essential to ensure that their personal information is kept secure and confidential.

- Ensuring data quality: ML models are only as good as the data they are trained on.
  Data policy and governance can help ensure that the data used for citizen projects are of high quality, accurate, and relevant.

- Promoting fairness and transparency: Bias and discrimination can creep into ML models if the data used to train them is biased.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

...

- Enabling responsible data use: ML models can have significant impacts on society, and it's essential to ensure that they are used in a responsible and ethical way.
- Facilitating collaboration and innovation: Data policy and governance can help establish a framework for collaboration between different stakeholders, including citizens, researchers, policymakers, and other stakeholders.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Conclusion

- Data is the backbone for ML in citizen projects to deliver significant benefits to society.

-However, it's essential to ensure that these projects are undertaken in a way that respects citizens' privacy and ensures responsible data use.

-A data policy and governance framework can help achieve these goals, and the recommendations are to be outlined.

-Ultimately, an effective data policy and governance framework can help ensure that citizen projects using ML are successful, beneficial, and respectful of citizens' rights.

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Reference

1 Amani Abou-Zeid. AU Data Policy Framework. Addis Ababa. 08, 84 (2022).

2 Abiy Ahmed ALi. Report. Digital Ethiopia 2025 – a Strategy for Ethiopia's Inclusive Prosperity. 2, 155 (2021).

3 Taffere Tesfachew. Report. Ethiopia's Drive to Advance Digital Transformation. 4.2, 155(2022). [4] EPHI Report. Guideline for Data Management and Sharing August Addis Ababa, Ethiopia. 2, 37 (2019)

Introduction
Motivation
Data access and sharing in Ethiopia
Data Management policy
Conclusion
Reference

ML for Citizen Science Projects

Thank you

Thank you